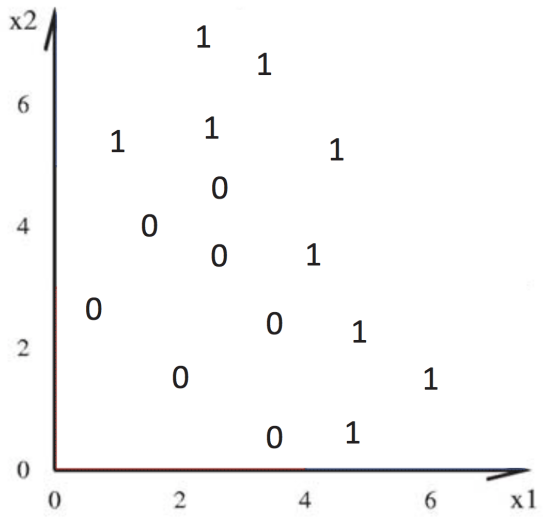
Intro to Decision Trees *(find and work with a partner)*

1. First, what is n (number of data points)? What is p (number of features)? Given the training data below and the decision tree shown on the slides, what is the classification error on this data?

Day	Outlook	Temperature	Humidity	Wind	PlayTennis (y)
x_1	Sunny	Hot	High	Weak	No
x_2	Sunny	Hot	High	Strong	No
x_3	Overcast	Hot	High	Weak	Yes
x_4	Rain	Mild	High	Weak	Yes
x_5	Rain	Cool	Normal	Weak	Yes
x_6	Rain	Cool	Normal	Strong	No
x_7	Overcast	Cool	Normal	Strong	Yes
x_8	Sunny	Mild	High	Weak	No
x_9	Sunny	Cool	Normal	Weak	Yes
x_{10}	Rain	Mild	Normal	Weak	Yes
x_{11}	Sunny	Mild	Normal	Strong	Yes
x_{12}	Overcast	Mild	High	Strong	Yes
x_{13}	Overcast	Hot	Normal	Weak	Yes
x_{14}	Rain	Mild	High	Strong	No

Data from Machine Learning by Tom Mitchell (Table 3.2)

2. Devise a decision tree for this data that perfectly classifies the given examples. (Here x_1 and x_2 refer to features of each example, so $p = 2$.) *Hint: start with the condition $x_2 < 3$ at the root.*



Based on example by Eric Eaton

3. Given the probabilities of each class below, what is the average number of bits needed to transmit one value?

Class	Probability	Encoding
Senior	0.5	
Junior	0.25	
Sophomore	0.125	
First-year	0.125	

Based on example from Ameet Soni

4. Given the movie data below (Liked (Li) is the response variable), select the feature the maximizes the information gain. This will be the root node in our decision tree.

Movie	Type	Length	Director	Famous actors	Liked?
m1	Comedy	Short	Adamson	No	Yes
m2	Animated	Short	Lasseter	No	No
m3	Drama	Medium	Adamson	No	Yes
m4	Animated	Long	Lasseter	Yes	No
m5	Comedy	Long	Lasseter	Yes	No
m6	Drama	Medium	Singer	Yes	Yes
m7	Animated	Short	Singer	No	Yes
m8	Comedy	Long	Adamson	Yes	Yes
m9	Drama	Medium	Lasseter	No	Yes

$P(\text{Li} = \text{yes}) = 2/3$
$H(\text{Li}) = 0.92$

$H(\text{Li} \mid \text{T}) = 0.61$
$H(\text{Li} \mid \text{Le}) = 0.61$
$H(\text{Li} \mid \text{D}) = 0.36$
$H(\text{Li} \mid \text{F}) = 0.85$

$\text{Gain}(\text{Li}, \text{T}) =$
$\text{Gain}(\text{Li}, \text{Le}) =$
$\text{Gain}(\text{Li}, \text{D}) =$
$\text{Gain}(\text{Li}, \text{F}) =$

Based on materials by Jessica Wu and Ziv Bar-Joseph